

DOCUMENT RESUME

ED 093 285

IR 000 766

AUTHOR Dailey, K. Anne
TITLE Compressed Speech: Potential Application for Air Force Technical Training. Final Report, August 73-November 73.
INSTITUTION Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO AFHRL-TR-74-27
PUB DATE Apr 74
NOTE 19p.
EDRS PRICE MF-\$0.75 HC-\$1.50 PLUS POSTAGE
DESCRIPTORS Instructional Technology; Language Research; Phonotape Recordings; Speech; *Speech Compression; *Technical Education; *Time Factors (Learning)
IDENTIFIERS *United States Air Force

ABSTRACT

Time-compressed speech (also called compressed speech, speeded speech, or accelerated speech) is an extension of the normal recording procedure for reproducing the spoken word. Compressed speech can be used to achieve dramatic reductions in listening time without significant loss in comprehension. The implications of such temporal reductions in the educational setting are that time savings can be used functionally to review, repeat, or extend the material which was originally presented. This report includes a definition and description of compressed speech techniques as they have been evolved, a discussion of certain methodological issues engendered by the current literature, a summary of selected literature (grouped by system topics), and a discussion of current and potential studies, possible applications, and general recommendations concerning the impact of compressed speech on technical training in the Air Force. (Author)

AIR FORCE



**HUMAN
RESOURCES**

**COMPRESSED SPEECH
POTENTIAL APPLICATION FOR AIR FORCE
TECHNICAL TRAINING**

By
K. Anne Dailey

**TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado 80230**

April 1974

Final Report for Period August 1973 - November 1973

Approved for public release; distribution unlimited.

LABORATORY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

**AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235**

ED 093285

IR 000 766

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Technical Training Division, Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado 80230, under project 1121, with Hq Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

MARTY. R. ROCKWAY, Technical Director
Technical Training Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TR-74-27	2. GOVT. ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPRESSED SPEECH: POTENTIAL APPLICATION FOR AIR FORCE TECHNICAL TRAINING		5. TYPE OF REPORT & PERIOD COVERED Final August 73 - November 73
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) K. Anne Dailey		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Technical Training Division Air Force Human Resources Laboratory Lowry Air Force Base, Colorado 80230		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6270F 11210113
11. CONTROLLING OFFICE NAME AND ADDRESS Hq Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		12. REPORT DATE April 1974
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 20
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES This technical report is an expansion of a presentation given at the 1973 convention of the Rocky Mountain Psychological Association, Las Vegas, Nevada, May 1973.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Compressed speech listening listening comprehension speeded speech time-compressed speech accelerated speech rate-controlled speech media comprehension speech compressors word rate		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Time-compressed speech (also denoted as compressed speech, speeded speech, or accelerated speech) is an extension of the normal recording procedure for reproducing the spoken word. Compressed speech can be used to achieve dramatic reductions in listening time without significant loss in comprehension. The implications of such temporal reductions in the educational setting are that time savings can be used functionally to review, repeat or extend the material which was originally presented. This report includes a definition and description of compressed speech techniques as they have been evolved, a discussion of certain methodological issues engendered by the current literature, a summary of selected literature (grouped by system topics), and a discussion of current and potential studies, possible applications, and general recommendations concerning the impact of compressed speech on		

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Block 20. (Continued)

technical training in the Air Force.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TABLE OF CONTENTS

	Page
I. Introduction	3
II. Definition and Description of Compressed Speech Techniques	3
III. Methodological Issues	6
IV. Summary of Literature on Compressed Speech	7
System Variables and Organismic Variables Which Affect Measures of Comprehension	7
V. Implications for Research	10
References	13
Appendix A. Speech Compression Facilities Available for Commercial Use	17

COMPRESSED SPEECH: POTENTIAL APPLICATION FOR AIR FORCE TECHNICAL TRAINING

I. INTRODUCTION

The time spent in listening to someone else speaking accounts for a major portion of daily information gathering, particularly in the classroom; yet speech may be one of the most inefficient methods for the communication of information, since it is comparatively slow and is subject to distraction. In an effort to devise methods for transmitting oral information more succinctly, investigators have examined some of the parameters of speech products which have been compressed in time.

When a tape of a spoken passage is played back at a rate which is faster than it was originally recorded, its word rate per minute will be increased or compressed. Such compression results in a proportionate rise in the pitch of the speaker's voice as well as in some distortion in the voice quality. If very brief segments (usually 30 to 40 milliseconds duration) are removed from the original recording and the resulting gaps are closed, the speed of playback can remain at its original pace, leaving the vocal pitch unaffected, but accelerating the word rate. It has been demonstrated that compressed speech can be used to achieve rather dramatic reductions in listening time without significant loss in comprehension. It may also be used as a method for assessing the critical features and channel capacities of cognitive or information-processing aspects related to the listening process in a manner which parallels the use of the tachistoscope in assessing the visual process.

It may be hypothesized, therefore, that effective use of compressed speech in the Air Force technical training environment should reduce the amount of time needed per unit of content material for the information transmitting process. Resultant savings in time might then be organized to facilitate such activities as presenting additional content material or for content enrichment through review or expansion of the material which was originally presented. It is suggested by this report that certain aspects of educational effectiveness and particularly efficiency of training might be enhanced by the functional exploration of these possibilities. Thus, this technical report will cover some of the germane issues concerning the present state-of-the-art of compressed speech; its content will, of course, be subject to update as new information concerning research and tech-

nology becomes available. Included are a definition and description of compressed speech techniques, a discussion of the methodological issues engendered by the current literature, a summary of the literature (grouped by system topics), and a discussion of current and potential studies, possible applications, and general recommendations concerning the impact of compressed speech on technical training in the Air Force.

II. DEFINITION AND DESCRIPTION OF COMPRESSED SPEECH TECHNIQUES

Time compressed speech (also denoted as compressed speech, speeded speech, or accelerated speech) is an extension of the normal recording procedure for reproducing the spoken word. When a spoken passage is transmitted at a rate that is faster than the normal speaking rate (approximately 125 to 175 words per minute), the time required for presentation is reduced and the message is referred to as being "compressed." Originally, research interest in speech compression was generated by one of the problems encountered in educating blind students. Reading in braille is a rather slow process so, as an alternative, educators began to use recorded oral presentations for transmitting information. Normal speech rates (125 to 175 wpm) present information at a rate which is close to $\frac{2}{3}$ of the normal reading rate of sighted adults (200 to 250 wpm) and $\frac{1}{2}$ the potential listening rate. Methods for compressing speech have evolved through an iterative process which has been dictated by changes in application demand and technological considerations.

Speed Changing Method. To increase the rate of information being transmitted orally so that it more closely approximated that of normal reading rates, some investigators (e.g., Klump & Webster, 1961) speeded up the tape recording. Although the change in tape speed did accomplish an increase in presentation rate, this increase also created distortions in voice pitch and quality which made the stimulus material unintelligible above certain speeds. However, from the very practical, applied need evidenced by blind students and the general dissatisfaction with the obvious technical limitations of the speed changing method, several other procedures for generating compressed speech were developed.

Sampling Methods. As a forerunner of compression techniques, a speech signal was demonstrated by Miller and Licklider (1950) to be intelligible even after large portions of the signal were removed through interrupting or "chopping" the signal by cutting it off and on with an electronic switch. Examining the independent variables of: (1) number of interruptions per second, (2) proportion of time the speech signal was on, and (3) degree of regularity of the interruptions, these investigators found that even with an interruption frequency of ten times per second, intelligibility did not drop below 90% of the base established with normal speech until 50% of the speech signal had been removed. However, Miller and Licklider's chopping technique did not close the resulting gaps between segments and therefore the speech was not compressed in time. For the future of compressed speech technology, the most important implication of this evidence was that the normal speech signal is redundant and seems to be rather resilient to having portions deleted.

The next logical step in developing the technology for producing a true time compressed signal without any frequency distortion was to remove segments from a tape recording and then abut the remaining segments. The first and rather primitive method for compressing speech signals was developed by Garvey and Henneman (1950) and Garvey (1953) who chopped and spliced the speech segments by hand, using a list of single spondaic (Words made up of two long syllables.) words rather than connected speech. The technology of this cumbersome method was described by Garvey as follows:

The position of the word on the tape was determined by slowly passing the tape over the magnetic playback head of the recorder. The beginning and ending of each word could be determined by auditory means and marked on the tape. Once the position of the words had been determined, the sections to be removed were marked on the tape with soft lead pencil and later cut out with shears. The ends of the segments left in were then spliced with scotch tape. If speedup of twice the original were desired, this could be accomplished by removing two centimeters, leaving two centimeters in, and so forth. (Garvey & Henneman, 1950, p.5)

The essential elements of this general technique, which has come to be called the sampling method, were that very brief segments (usually 30 to 40 milliseconds) of the original recording were cut out and discarded and the resulting gaps were closed. These discarded segments were so small that not even the shortest speech sounds were excised and their absence generally could not be

detected. The speed of playback was not changed, and therefore, the vocal pitch remained the same, while the word rate was increased. However, since this method required splicing the tape manually, it was so tedious and time-consuming that utilizing this method for other than purely experimental purposes was not practical.

Not long after the publication of Garvey's work, an electromechanical speech compressor was demonstrated at the University of Illinois (Fairbanks, Everitt, & Jaeger, 1954). This compressing device is more a tape reproducer than a tape recorder. It employs the sampling method by having a movable cylinder with four recording heads which are equally spaced around the circumference of the cylinder. The amount of compression is a function of the tape speed and the rate of cylinder rotation. A signal, which is identical to that on the original tape, may be recorded by having the cylinder remain stationary and the *prerecorded* tape move around one recording head on the cylinder at normal recording speed. Compressed rates of speech are generated by rotating the cylinder. The discard interval is equal to the separation along the circumference between adjacent heads, divided by the tape speed. The most lucid description of the operation of Fairbanks' device can be found in Van Matre and Steinemann (1972) who write that, when a particular degree of compression is required,

the tape speed is increased and the cylinder is rotated in the direction of the tape motion at a speed which maintains the normal tape speed as the tape passes around the cylinder. By keeping the relative speed of the tape as it passes across the four recording heads equal to the normal recording speed, voice pitch is maintained at a normal level. Under these conditions, each tape head briefly makes contact with the tape and reproduces, as originally recorded, the material on only that portion of the tape. The final recording then consists of only those samples of the original tape which made contact with the four moving tape heads. With this type of sampling technique, the tape portions between the heads are randomly discarded. (Van Matre & Steinemann, 1972).

Conversely, speech which is "expanded" can be produced on this device by running the original tape at a slower than normal speed and rotating the cylinder in the reverse direction. The Fairbanks method is ingenious, but suffers from the use of high speed rotating mechanical components and difficulty with signal amplitude modulation due to mistaching of signal output from the four heads. A revised version of the Fairbanks' device is presently being manufactured and

distributed under the name of Graham Compressor or "Whirling Dervish."

Sampling methods which utilize computers to discard speech segments on the basis of programmed rules are, not surprisingly, called computer sampling methods. When very short segments of speech are stored in the computer, the computer can be programmed to discard specified samples and consecutively reproduce the remaining segments. Unlike the Fairbanks' method, the computer can sample on the basis of some rule which is other than periodic. For example, some speech sounds are so short that intelligibility may be preserved by leaving them intact, whereas other speech sounds are so long in duration that intelligibility may be unaffected by discarding proportionately larger segments. A computer program can make such distinctions with ease. (Ball, 1961, developed the first known computer program for this purpose, followed by programs developed by Scott, 1965, and Cramer and Talambiras, 1970). Although seemingly attractive as both a production and a research tool, this method has not been utilized sufficiently to ascertain fully its flexibility because the cost of using the computer has been somewhat prohibitive.

Bell Telephone Company has developed a similar technique, called Time-Assigned-Speech-Interpolation (TASI), which increases the efficiency of telephone line utilization through computerized electromechanical switching; i.e., sampling of oral signals. Currently in use with undersea telephone technology, computers check for pauses in conversations by sampling each telephone channel 8,000 times a second. When a pause is detected, TASI very effectively multiplies the number of channels available by transmitting other conversations on the line. Although the TASI sampling method is not a periodic sampling procedure, since it depends upon the presence or absence of a speech signal, the principle of increasing transmission efficiency by interweaving signals during speech pauses is analogous to the general principle of compressed speech.

Harmonic Compression. The American Foundation for the Blind constructed the first harmonic compressor from plans and specifications made available by Bell Laboratories. Harmonic compression achieves its results by frequency division. Through a filtering system, the harmonic compressor separates the component frequencies in a recorded signal of speech. These component frequencies are then reduced by one-half and mixed together again. The resulting speech is one

octave lower than the original speech and, therefore, can be reproduced at twice the recording speed. The result of this method is that the pitch of the speaker's voice is restored to its original level and the speed of the presentation is compressed to 50% of its original production time. With this method, no advantage is taken of excising pause time or portions of long vowel sounds to achieve speech efficiency, but since no excising of speech segments is made either, all phonetic information remains intact. The most severe drawback to the technique is that this is not a variable speed procedure, since the only compression rate utilized is twice as fast as the original. In principle, harmonic compression could reduce the component frequencies by *any* constant, not just by one-half, but this possibility does not seem to have been explored.

Current Technology. Two newer techniques which have become commercially available are (1) a *voice activator recorder* or *pause compressor* which excises the pauses between words and sentences and (2) an *electronic pitch signal processor* which converts the voice signal into a digital format, compresses the signal, then converts it back into an analog or voice representation. The voice activator does not distort the phonetic information, but does distort linguistic pacing by eliminating the juncture pauses between sentences, and thus may adversely affect comprehension. In addition, the high-speed stop-start clutch system is mechanically fallible. The pitch-signal processor is a small digital computer in conjunction with a quality tape deck which is highly reliable. This sampling method reduces all sounds and pauses by the same percentage, which preserves the relative rhythm of the original recording, but at very high compression rates may lose some phonetic information. The point at which any loss in comprehension becomes critical due to the specific features of these machines has not yet been determined in the literature and so far no direct comparison has been made between the two methods.

In summary, four methods for accelerating the speed of speech have been described. The *speed changing method* alters word rate by changing the rate at which the tape passes over the tape head without performing any other functions on the tape, but its usefulness is severely limited because it proportionately distorts the pitch and quality of the speaker's voice. The *sampling method*, either by *electromechanical* compression or by *computer*, changes the word rate of recorded speech by almost any desired amount without

distorting the pitch or quality of the speaker's voice by excising very small segments from a recorded tape. The *harmonic compressor* compresses the word rate without distortion in vocal pitch or quality by frequency division and then restoration with temporal alteration, but has only been used to double the word rate. The *pause compressor* leaves the phonetic information unchanged, but seriously distorts the linguistic pacing which may affect comprehension. Each of these general methods for compressing speech represents a technological breakthrough for the field and has its own important features and limitations. New devices and improvements on older procedures are being developed currently, but since the "ultimate" mechanism which can compress speech over a large range of wpm rates at a reasonable cost is not yet available, potential investigators are urged to take advantage of one of the speech compression facilities which are available for commercial use (see Appendix A for listing), rather than to invest in an apparatus which may prove to be part of the developmental stage of the art of compression.

III. METHODOLOGICAL ISSUES

As in other areas of psychology and education, certain substantive methodological dissimilarities plague the reporting and interpretation of compressed speech literature. The most difficult issue to take into account is that of estimating and/or comparing reported rates of compression. It should be emphasized here that the final compressed wpm rate or percentage of time saved is very much dependent upon the original wpm rate. Four indices are in common use, but none has a linear relationship with any other. Some investigators report only the *percentage of time saved* (e.g., 50% time-saving) without reporting either the original uncompressed or the final compressed word rate. This index can be roughly translated into the second index — reporting findings in terms of *accelerative speed* (e.g., 50% acceleration). However, if the original word rate is not reported, then again, the measure is relatively difficult to assess. *Word rate* is the most frequently reported index and through general use has become probably the most functional. However, two substantial problems about relying on this index are the insensitivity of this measure to differences in stimulus difficulty and the lack of uniform methods for calculating word rate. Evaluation problems are compounded by the fact that researchers seldom make any statement about

how the index was achieved (e.g., using the actual word count divided by the time involved or using estimating formulas to achieve word counts). The most accurate, but least frequently used, method is to report the *syllable rate*. This method is sensitive to various difficulty levels of material and, therefore, is a potentially powerful index in terms of the amount of information it provides. It is not, however, linearly related to word rate (which appears to be becoming the standard) and is, therefore, not very useful when trying to compare various results already in the literature.

Just as problematical to evaluating and comparing results is the reporting of dependent measures with unknown psychometric properties. Measurement of change is, as usual, a subtle and often tricky enterprise and is as important in assessing performance in studies of compressed speech as elsewhere. The usual procedure for measuring comprehension of a passage is to administer a multiple-choice test directly following the presentation of the passage at a particular compression rate. Evidence regarding the effect of compression rate upon comprehension is gathered by comparing scores achieved by experimental groups which have heard the message at different rates.

The most difficult problem is to produce a research instrument which is a psychometrically adequate test by having demonstrated indices of reliability and validity. Nowhere in the literature on compressed speech is there an indication that the investigators took psychometric precautions with their measurement tools. Secondly, any test of comprehension ought to be an unbiased and representative sample from the domain of material to be assessed. Again, little indication is given by researchers that any effort was made to validate their instruments through domain sampling. Thirdly, it is difficult to assess the effects of prior knowledge on comprehension scores unless the experimenter uses textual material which has a low exposure probability or uses control groups to assess base knowledge before exposure to the experimental material. No evidence of sensitivity to this issue was found to be demonstrated in the literature.

One study, which utilized a radically different methodology for measuring comprehension (Downing, 1971), solved the problem of domain sampling, but not the problem of prior knowledge. The technique is called the "cloze procedure" and evolved as a method for measuring the readability of written passages. It consists of periodic (or, in

some cases, random) deletion of words within a given passage. The deleted words are replaced with blank spaces and the subject's task is to guess the missing words. His comprehension score is the total number of correct guesses. Downing adapted the cloze procedure to measure oral speech and found that this procedure produced results which were very similar to the multiple-choice test results. However, Carver (1973) found that neither of the two types of cloze measures which he examined were very sensitive to the different presentation rates (75 to 450 wpm) since the cloze test scores changed only 20% over the rates sampled while other measures of understanding changed 80%. Additionally, a serious drawback for using the cloze procedure as a technique for measuring oral comprehension is the requirement for hand scoring of the items, since the technique requires that the subject must write down his answers.

IV. SUMMARY OF LITERATURE ON COMPRESSED SPEECH

System Variables and Organismic Variables Which Affect Measures of Comprehension

The following is a review of the literature on compressed speech emphasizing those studies which appear to be most relevant to potential application to Air Force technical training situations. The discussion is divided into two subsections: system variables, and organismic variables. System variables are those factors which are part of the apparatus or procedure of the research paradigm; organismic variables are those factors which are examined for their possible interactive effect with system variables and which differentially affect measures of comprehension. Word rate, method of compression or presentation, and difficulty of material are discussed as system variables; mental ability, practice or pretraining, reading rate, and long term retention are discussed as organismic variables. The dependent measure for all of the studies under examination is some kind of multiple-choice test of comprehension. Since it would not elucidate the present application, no coverage of the literature concerning blind subjects or elementary school subjects is given.

System Variable: Word Rate. The first attempts to vary the number of words per minute of an oral presentation used trained speakers who systematically increased the pace at which they spoke

and produced rates ranging from 100 to 285 words per minute (wpm). The results of rapid speaking are not the same as altering the rate on a compressor, however, since both deterioration in articulation and reduction in pause time occur, and in the upper limits the speaker begins to approach the truncated version of the speech an auctioneer uses. Three early examinations of increased word rates produced by trained speakers found no significant decreases in comprehension scores up to 225 wpm (Goldstein, 1940; Harwood, 1955; Nelson, 1948).

Another technical variation for producing accelerated word rates was reported by Diehl, White, and Burk (1959). These investigators excised only portions of the pause time between words to achieve accelerated speech rates of 160 wpm (50% of pause time removed) and 192 wpm (75% of pause time removed). No significant difference in comprehension scores was found between the control group who experienced normal speech and either of the experimental groups. Nor, notably, were there differences in the subjective ratings of the speaker's delivery. Since this method is limited as to how much the rate can be accelerated merely through excising pauses (i.e., 200 wpm is the approximate upper limit), the methodology has not been pursued until recently, when it has been utilized with some success in combination with an electronic sampling method which also shortens the vowel type sounds.

The paradigm for most recent studies of the effect of word rate on comprehension can be generalized from the work of Emerson Foulke who established a Center for Rate Controlled Recordings at the University of Louisville in Kentucky. Besides conducting research and providing compressed recordings, this Center serves as a clearinghouse and publishes a newsletter for contemporary literature concerning compressed speech research. Generally, the following factors can describe the basic parameters of the experimentation. Foulke and his associates (Foulke, 1966; Foulke, 1968a; 1968b; Foulke & Robinson, 1969) use: (1) a Fairbanks sampling compressor to generate the stimulus material; (2) rates ranging from 125 to as much as 400 wpm in increments of 25 or 50 wpm; (3) different groups of subjects to test different rates of speech (no repeated measures on subjects); and (4) no control group is utilized; e.g., for testing subjects' comprehension without listening to the stimulus passage and, therefore whether comprehension scores at high rates of compression are significantly different from zero, cannot be determined.

Examinations of the effect of word rate on dependent measures of comprehension corroborate that no significant decline in comprehension scores occurs up to 275 wpm and severe declines occur only above 350 wpm (Downing, 1971; Fairbanks & Kodman, 1958; Garvey, 1953; Friedman, Orr, Freedle, & Norris, 1966; Langford, 1968). Foulke (1968) examined rates from 124 to 400 wpm by increments of 25 wpm and found that comprehension did not decline at a statistically significant level until 325 wpm; Fairbanks, Guttman, and Miron (1957) found no significant loss of comprehension at 141, 207, 282 wpm, but reported a significant loss at 470 wpm; Foulke, Amster, Nolan, and Bixler (1962) found no significant loss up to 275 wpm, but a significant loss in comprehension between 275 and 375 wpm. One observation which may be raised, however, is that in some studies, more data points might have been sampled, since it seems experimentally weak to sample at 275 wpm and then again only at very high rates of 375 or 470 wpm.

It appears that the rate at which comprehension declines as word rate increases is not a simple linear function, but some complex function approaching an exponential relationship above 350 wpm. However, for technical training situations, rates above 300 wpm do not seem to warrant pragmatic application since, functionally, 300 wpm is almost twice as rapid a transmission rate as is normally used in the lecture mode (150 to 175 wpm) and the savings in time are already substantial. Above 300 wpm, the effort to attend seems to become too taxing and concentration suffers. Of course, factors contributing to any congenital limitations in human channel capacities may be attributed to lack of training in how to listen efficiently and/or to sound production system defects which impede listenability (such as the "glottle stop" sound of high compression). If a cleaner production technique could be achieved and training in how to listen given, the ceiling for comprehending compressed speech might be raised.

System Variable: Method of Compression or Presentation. Objective measures of comprehension and subjective measures of listener's preference both indicate that the electromechanical sampling method of compression is more effective than the method of increasing tape speed (Foulke & Sticht, 1969). Presumably, the lack of effectiveness of increasing tape speed is some function of the resulting pitch and vocal quality distortion. Although the speech quality alterations produced by electromechanical sampling are perceptible and vary with the degree of compression, these alter-

ations do not seem to be perceived as being as annoying as the "Donald Duck" quality of the speed changing method.

Modes for hearing compressed speech were investigated in a comparison of listening to compressed speech using headphones versus a public address system (George, 1970). This study reported no significant differences in comprehension as a function of mode of presentation and found no significant interaction between mode of presentation and compression rate. Such a finding implies that presentation of compressed speech to large audiences may be successfully accomplished.

At a college learning center Tonra (1972), developed a multimedia self-paced course, designed to include as many learning options as possible to allow the student to choose when and how he would learn. This introductory psychology course had weekly lectures available in either normal or compressed form and Scientific American articles made available in five presentation modes: (1) reprint of article; (2) tape of reprint at normal speed; (3) tape and reprint together (normal speed); (4) compressed tape of reprint (25 to 35% compression); and (5) compressed tape and reprint together. Using the dependent variables of record keeping and a questionnaire, Tonra reported that student preferences were for the reprint itself followed by 1/3 of the students (from 40 to 60, depending upon enrollment) consistently choosing the compressed tape and reading option of presentation mode number 5. In addition, after the introduction of the compressed tapes of lectures, there was a 50% increase in use of lectures in the Learning Center and the "comments" section of the questionnaire indicated that the students used the compressed lectures primarily as an organizing and review tool, rather than solely as an information source. This study confirms that, given a choice, students will utilize compressed speech as a useful resource.

Gerber (1968), acting on an idea suggested by Scott (1965), investigated the effects of dichotic versus diotic presentation of speeded speech. Diotic presentation is the normal mode of presentation in which the same stimulus material (e.g., generated by the Fairbanks Sampling Method) is presented to both ears; whereas in a dichotic presentation the normally discarded segments are recovered, another tape of these segments is made up so that they occur temporally at the same point as the compressed word from which they were excised, and the formerly discarded segments are

played for one ear simultaneously with the compressed version for the other ear. Gerber found the dichotic presentation to be more intelligible than the diotic presentation and suggested that this method is not only feasible, but preferable for transmitting information. In a follow-up study, Gerber and Scott (1970) found the average intelligibility exceeded 97% for the aggregate of dichotic signals, while the average intelligibility for the aggregate of diotic signals was just over 93%. This difference was reported to be significant beyond the .01 level. Substantiating these results, Brown (1972), using both normal and 2:1 compression in a study of unilateral (diotic) and dichotic variation in presentation mode, reported no significant differences between measures of comprehension for dichotic normal and dichotic 2:1 compressed modes.

These last three studies raise further questions about the technology of compressing speech. Is the most significant factor in understanding compressed speech the duration of the discarded interval or the duration of the remaining interval? Are there differences existing for the optimal length of the discard interval for dichotic and for diotic presentations? However, considering the additional costs of preparing dichotic material, it seems to the present author that even though the differences between diotic and dichotic modes of presentation were found to be statistically significant, the differences were too small to be of much practical significance to the implementation of any large scale program which utilizes compressed speech.

System Variable: Difficulty of Material. Several investigations of the interaction effect of difficulty of material and compression rate on measures of comprehension have been conducted. Generally, these studies have presented literary and scientific material as the easy and hard stimuli and all but one of the studies examined found no evidence for comprehension of difficult or easy material being differentially affected as a function of compression rate (Foulke, Amster, Nolan, & Bixler, 1962; Fairbanks, Guttman, & Miron, 1957; Goldhaber & Weaver, 1968; Goldstein, 1940; Harwood, 1955; Nelson, 1948; Reid, 1968).

Only Bixler, Foulke, Amster, and Nolan (1961) found a statistically significant interaction between rate and type of material in a study which presented both literary and scientific materials to blind subjects at word rates of 175, 225, 275, 325, and 375 wpm. However, it was concluded by the authors that this interaction was due to the

reduced range within which the initially lower scores on the scientific passage could vary. The reduced range of comprehension scores for the most difficult selections indicates other procedural problems which exist in designing and interpreting studies of this variable. It is difficult to assess the motivational contribution from differential intrinsic interest in stimulus material of such disparate difficulty. Fundamentally, no assessment procedure has been developed which can successfully quantify the differences in level of difficulty between particular selections.

Organismic Variable: Mental Ability. Understandably, the literature supports a positive relationship between listening comprehension and measurements of mental ability, but comprehension scores are not found to be differentially affected as a function of any mental ability by word rate interaction (Downing, 1971; Fairbanks, Guttman, & Miron, 1957; Goldstein, 1940; Lanford, 1968; Michel-Miller, 1970; Nelson, 1948; Sticht, 1968, 1969, 1970, 1971). As representative of the mental ability studies, Sticht (1969) presented passages at 175, 275, and 424 wpm to Army recruits of three levels of mental ability as measured by the AFQT (10 to 30, 31 to 64, and 65 to 100). For all three ability groups, there was only a slight decline in comprehension from 175 to 275 wpm and a more rapid decline at 425 wpm. This finding is consistent with the literature, but for more compelling generalizability additional points should have been sampled on the compressed word rate continuum. The main point to be stressed is that again, no significant interaction was found between compression rate and aptitude.

When a significant interaction is found (Parker, 1969), it is difficult to interpret. Comprehension scores for the higher ability groups are initially larger and, therefore, have a wider range in which to vary than the lower ability groups, whose scores are initially closer to chance level and, therefore, remain close to chance level over the range of increased compression rates. Such non-homogeneity of variances may mask a main effect, while creating a spurious interaction effect.

Organismic Variable: Effects of Practice or Pretraining. The issue of whether practice of pretraining has an effect on comprehension scores as a function of compression rate has not been clarified. Foulke (1964) found no significant practice effects, whereas, Voor and Miller (1965) and Orr and Friedman, (1968) and Orr, Friedman, and Williams (1964) reported improvement with practice. An interaction effect, which suggested

that for easy material, previous exposure may not produce an increase in comprehension, while for difficult material, previous exposures may facilitate comprehension, was found by Reid (1968). The issue needs to be resolved since it could have serious import for any sustained program using compressed speech as a training media.

Another facet of this area, which certainly may have an effect on stabilizing performance responses as well as motivational implications, is the issue of acclimation to the level of compression. Compressed speech is an unusual medium which has properties that have not been encountered in the normal course of the educational process. As such, the additional psychological requirement of active attending should also be examined systematically.

Organismic Variable: Reading Rate. Thames and Rossiter (1972) investigated the use of compressed speech as a reading rate pacer with high school sophomores. A control group read 10 stories on successive days while an experimental group read the same 10 stories while simultaneously listening to a recording of the stories which was played to pace the reading rate. The recording was played at a rate of 150 wpm on the first two days; thereafter, the wpm rate was increased by 50 wpm each succeeding day until the experimental group read on the last two days paced by a recording which had been compressed to the rate of 350 wpm. No significant differences were found between the control and experimental group in terms of comprehension, but the reading rate of the experimental group increased significantly over that of the control group when measured immediately after the experimental procedure. Unfortunately, no long term effects were examined in this study, so it is not possible to ascertain whether experience with pacing had any lasting effects on reading rates.

Organismic Variable: Long Term Retention as a Dependent Measure. Only a few studies have investigated long term retention of compressed speech stimuli. Of the seven examinations in the literature which use both long term retention and more immediate comprehension tests as dependent measures of performance, all found that retention was not adversely affected by accelerated speech either as a main effect or as an interaction effect with compression rate (Barabasz, 1968; Foulke, 1966, 1968; Foulke & Robinson, 1969; Short, 1971; Tonra, 1972; Woodcock & Clark, 1968). Since long term retention is of major importance to the goals of any training situation, these reports are encouraging for the development of both research and practical applications of compressed speech.

In summary, the research findings described in this paper can be recapitulated as follows: (1) There are clear indications that comprehension is only nominally affected by accelerated word rates up to twice the normal pace. (2) Considerable work needs to be done to develop assessment procedures for describing the differences between difficult and easy material and the whole area of measurement devices and techniques needs examination. (4) For technical training applications, the several findings of no significant interaction between mental ability and word rate as reflected in comprehension scores are encouraging as they imply that no special considerations need to be made for the Air Force student population of varying abilities. (5) Since the literature is inconclusive, the effects of practice or pretraining still need to be determined empirically.

Considering the weaknesses and strengths demonstrated by the literature, the following is a list of experimental procedures which might be applied when setting up design parameters for studies of compressed speech:

- (1) make an effort to obtain assessment instruments with reasonable psychometric properties;
- (2) use as many marker variables or predictors as are possible within the time and cost constraints in order to be able to make generalizations to the large bodies of extant information;
- (3) obtain base line rates by administering measurement instruments to naive subjects who do not listen to the experimental selections (i.e., employ control groups);
- (4) sample systematically from the compression rate continuum so that there are no large gaps which make word rate functions difficult to interpret; and
- (5) investigate the feasibility of using within subject repeated measures designs in order to control for variance due to individual differences (e.g., ability, reading rate, etc.) and still be parsimonious with subject resources.

V. IMPLICATIONS FOR RESEARCH

This report has described the relevant elements of the technology of compressed speech, given a delineation of compressed speech measurement problems, and presented a description of the functional relationships which appear to exist between measures of performance with compressed speech and impacting system and organismic variables. Based on the foregoing

considerations, the following are suggestions for feasible research approaches toward development of practical compressed speech research programs with potential operational payoff to the Air Force in the near future. The rationale for embarking on a formal program of research and evaluation of compressed speech, as a medium for use in Air Force Technical Training situations, is developed within the context of the following particular advantages which might accrue: (1) used at a moderate level in primary instruction, compressed speech could save instruction time without adversely affecting comprehension; (2) compressed speech could prove to be a powerful and efficient tool for reviewing material and for remedial work; (3) with the accompaniment of some formal training, compressed speech could increase and improve listening skills; (4) as a novel medium, compressed speech could improve motivation through increased curriculum variety; and (5) as a diagnostic or methodological tool, compressed speech might be used to assess cognitive functions and difficulties.

In view of the fact that certain parameters of compressed speech have been fairly well documented, the feasibility and limitations of applying this medium within the context of requirements for technical training situations should be determined. The need for such determination suggests a concerted research program which can focus on exploiting the known advantages, while at the same time systematically examining the unclear areas. The following is a description of some of the efforts in which the Air Force is currently involved which may answer critical and relevant questions.

1. The Learning Center at the Air University (Personal communication with Lt Col William P. Pannell.) has an ongoing program which has made 31 hours of the lecture material used in the Academic Instructor and Allied Officer School available in compressed tape/slide format. This material is being evaluated for its usefulness as a medium to facilitate make-up work (e.g., when student misses class due to illness, etc.), for remedial work, for review purposes, and for primary instruction. Since the lectures are available in both normally paced and compressed form, evaluations can be made by noting if and how subjects utilize the compressed version and comparisons can be made between the two modes by utilizing normal classroom performance indicators.

2. The Learning Center Support Division at Randolph AFB, Texas (As reported by Captain Marlin J. Schaefer in a personal communication.) reproduced ten sound-slide programs at two compression rates (40% and 60% of original rate). The only data collected were the students' observations on the ATC Form 895, Learning Center Program Improvement Sheets. The response was overwhelmingly in favor of the 40% compression rate. The students stated the compressed versions made them feel more alert, cut their viewing time by approximately one-half, and seemed to increase their comprehension. The basic response to the 60% compression rate was that it was too fast, and did not allow the students enough time to concentrate on the slides and listen to the narration. On the basis of the students' comments, a speech compression machine has been ordered and the Learning Center plans to compress all of the sound-slide tapes at a moderate rate (30 to 40% compression) and give the student the option of listening to the normal or compressed version.

3. The Air Force Human Resources Laboratory is in the process of working in a consulting relationship with the Air University in setting up a study at the Learning Center at the Air University which will examine temporal constraints which might impinge on the effective long term use of compressed speech. This study would:

(a) determine how long presentations can be before the effort of attending becomes too taxing;

(b) determine how often presentations can be given and still be effective (e.g., examining the effects of contrast, novelty, motivation, saturation); and

(c) determine impact for 8-hour training day (e.g., if information is compressed into a shorter length of time, can the student absorb additional amounts of information or are there information saturation ceilings?)

An example of such an experimental paradigm might present the compressed stimuli for stretches of three minutes, then allow a one minute respite, during which the student interacts with a prepared work sheet to answer questions which would consolidate the material which has just been covered. Three minutes of compressed speech would be alternated with one minute of worksheet or rest until performance measures indicate that

attention has become markedly affected. This mode of presentation would be compared to a format which presents a continuous compressed version of the material.

4. The Air Force Human Resources Laboratory is also in the process of trying to obtain a compression machine which is capable of doing the production work for a tape/slide presentation for each of five different In Service Instructor Training Courses. Since these courses vary along a continuum from general orientation low-content to highly specialized training, this study will help to determine what kinds of material are most effective for presentation in compressed mode. In addition to having the material made available at the normal pacing and at a compressed rate, the Air Force Human Resources Laboratory hopes to be able to give the subject variable speed capability and urge him to seek his own optimal level of compression. This latter capacity will allow comparison of the subject's preferred compressed rate with his reading rate for various types of content material.

Further examination of the relationship between reading rates and listening rates can be conducted by: (a) establishing individual reading speeds for particular kinds of material; (b) preparing verbal material to be heard at that optimal speed (i.e., match auditory rate with visual rate); and (c) comparing the performance effectiveness of matched compression with that of normal auditory rates to determine an optimality function.

It seems apparent from the documented research that a moderate degree of speech compression can not only speed information transmission, but result in a negligible loss in performance measures of comprehension. Considering the amount of time students spend in listening, if some of the temporal difference between thought rate and speech rate were eliminated by increasing the rate of speech, substantial payoffs in terms of a fuller maximization of communication efficiency than presently exists might be achieved.

REFERENCES

- Ball, J.H. Synthesis of original vocal pitch in accelerated playback. Unpublished master's thesis, Massachusetts Institute of Technology, 1961.
- Barabasz, A.F. A study of recall and retention of accelerated lecture presentation. *The Journal of Communication*, 1968, 18, 283-287.
- Bixler, R.H., Foulke, E., Amster, C., & Nolan, C. *Comprehension of rapid speech by the blind. Part I*. HEW Progress Report No. 1005, 1961.
- Brown, E.R. Investigation of comprehension of compressed and non-compressed speech with unilateral and dichotic variations. *CRCR Newsletter*, June 1972, 6 (6).
- Carver, R.P. Understanding, information-processing, and learning from prose materials. *Journal of Educational Psychology*, 1973, in press.
- Cramer, H.L., & Talambras, R. The braided speech method of time compressing speech. *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.
- Diehl, C.F., White, R.C., & Burk, K. Rate and communication. *Speech Monographs*, 1959, 26, 229-232.
- Downing, D. I. *The effects of speech rate, listener aptitude, and passage listenability upon listening comprehension*. (Doctoral Dissertation, Purdue University) Lafayette, Indiana, 1971.
- Fairbanks, G., Everitt, W.L., & Jaeger, R.P. Method for time or frequency compression-expansion of speech. *Transactions of the Institute of Radio Engineers, Professional Group on Audio*, 1954, AU-2, 7-12. Also in *Convention Record of the Institute of Radio Engineers, 1953 National Convention, Part 8: Information Theory*, 1953, 120-124.
- Fairbanks, G., Guttman, N., & Miron, M.S. Effects of time compression upon the comprehension of connected speech. *Journal of Speech and Hearing Disorders*, 1957, 22, 10-19.
- Fairbanks, G., & Kodman, F., Jr. Word intelligibility as a function of time compression. *Journal of the Acoustical Society of America*, 1958, 29, 636-641.
- Foulke, E. *The comprehension of rapid speech by the blind: Part III*. Final Progress Report, Project No. 2430, March 1964-June 1968.(a)
- Foulke, E. Listening comprehension as a function of a word rate. *The Journal of Communication*, 1968, 18, 198-206.(b)
- Foulke, E. The retention of information presented at an accelerated word rate. *The International Journal for the Education of the Blind*, October 1966.
- Foulke, E., Amster, C.H., Nolan, C.Y., & Bixler, R.H. The comprehension of rapid speech by the blind. *Exceptional Children*, 1962, 29, 134-141.
- Foulke, E., & Robinson, J. *The development of accelerated speech as a useful communication tool in the education of blind and other handicapped children*. Progress Report, Office of Education, HEW, Project No. 7-1254, June 1969.
- Foulke, E., & Sticht, T. Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin*, 1969, 72(1), 50-62.

- Friedman, L., Orr, D.B., Freedle, R.O., & Norris, C.M. Further research on speeded speech as an educational medium. *Progress Report No. 2*. Grant No. 7-48-7670-267. US Office of Education. American Institutes for Research, Washington Offices Communication Research Program, January 1966. (AIR-E-50-1/66-TR(2).)
- Garvey, W.D. The intelligibility of speeded speech. *Quarterly Journal of Speech*, 1953, **39**, 296-306.
- Garvey, W.D., & Henneman, R.H. Practical limits of speeded speech. United States Air Force Technical Report No. 5917. Wright-Patterson Air Force Base, Ohio, 1950.
- George, R.G. *Comprehension of compressed speech: Headsets versus public address system*. In CRCR Newsletter, April 15, 1970, Vol. 4, No. 4.
- Gerber, S.E. Dichotic and diotic presentation of speeded speech. *The Journal of Communication*, September 1968, **18**, 272-282.
- Gerber, S.E., & Scott, R.J. Dichotic speech-time compression. *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*. Louisville: University of Louisville, Perceptual Alternatives Laboratory, 1970.
- Goldhaber, G.M., & Weaver, C.H. Listener comprehension of compressed speech when the difficulty, rate of presentation, and sex of the listener are varied. *Speech Monographs*, 1968, **35**, 20-25.
- Goldstein, H. *Reading and listening comprehension at various controlled rates*. Teachers College, Columbia University Contributions to Education, No. 821. New York: Bureau of Publications, Teachers College, Columbia University, 1940.
- Harwood, K.A. Listenability and rate of presentation. *Speech Monographs*, 1955, **22**, 57-59.
- Klump, R.G., & Webster, J.C. Intelligibility of time-compressed speech. *Journal of the Acoustical Society of America*, 1961, **31**, 265-267.
- Langford, R.P. The effect of time compressed speech on listening comprehension. (Doctoral Dissertation, New York University) New York, N.Y., 1968.
- Michel-Miller, L. Levels of comprehension and activation under three conditions of time-compressed speech. Master's thesis, Florida State University, 1970.
- Miller, G.A., & Licklider, J.C.R. The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 1950, **22**, 167-173.
- Nelson, H.E. The effect of variations in rate on recall by radio listeners of "straight" newscasts. Summary: *Speech Monographs*, 1948, **15**, 173-80.
- Orr, D.B., & Friedman, H.L. Effect of massed practice on the comprehension of time-compressed speech. *Journal of Educational Psychology*, 1968, **59**, 6-41.
- Orr, D.B., Friedman, H.L., & Williams, J.C. Trainability of listening comprehension of speeded discourse. *Journal of Educational Psychology*, 1964, **56**, 148-156.
- Parker, C.C. Effect of rate of compression and mode of presentation on the comprehension of a recorded communication to junior college students of varying aptitudes. Paper presented to the Louisville Conference on Time Compressed Speech, October 1969.
- Reid, R.H. Grammatical complexity and comprehension of compressed speech. *The Journal of Communication*, 1968, **18**, 236-242.

- Scott, R.J.** Temporal effects in speech analysis and synthesis. Unpublished doctoral dissertation, University of Michigan, 1965. Also in Communication Scientist Laboratory, Contract No. NONR 1224 (22), NR 049-122, 1965.
- Short, S.H.** *Compressed speech tapes used with self-instruction course.* In CRCR Newsletter, February 15, 1971, Vol. 5, No. 2.
- Sticht, T.G.** Failure to increase learning using the time saved by the time compression of speech. *Journal of Educational Psychology*, 1971, **62**, 55-59.
- Sticht, T.G.** *Learning by listening in relation to aptitude reading, and rate-controlled speech.* Technical Report 69-23. HumRRO, Division No. 3, December 1969.
- Sticht, T.G.** Mental aptitude and comprehension of time-compressed and compressed-expanded listening selections. *Journal of Auditory Research*, 1970, **10**, 103-109.
- Sticht, T.G.** Some relationships of mental aptitude, reading ability, and listening ability using normal and time-compressed speech. *The Journal of Communication*, 1968, **18**, 243-258.
- Thomas, K.H., & Rossiter, C.M., Jr.** The effects of reading practice with compressed speech on reading rate and listening comprehension. *Audio Visual Communications Review*, 1972, **20**, 35-42.
- Tonra, B.** *Compressed speech used in introductory course.* In CRCR Newsletter, April 15, 1972, Vol. 6, No. 4.
- Van Matre, N.H., & Steinemann, J.H.** *A selective review of listening research.* Technical Bulletin STB 72-10, Naval Personnel and Training Research Laboratory, San Diego, California, May 1972.
- Voor, J.B., & Miller, J.M.** The effect of practice on the comprehension of speeded speech. *Speech Monographs*, 1965, **32**, 452-255.
- Woodcock, R.W., & Clark, C.R.** Comprehension of a narrative passage by elementary school children as a function of listening rate, rate, retention period, and IQ. *The Journal of Communication*, 1968, **18**, 259-271.

APPENDIX A

SPEECH COMPRESSION FACILITIES AVAILABLE FOR COMMERCIAL USE

Center for Instructional Technology
Wayne State University
5035 Woodward Avenue
Detroit, Michigan 48202
CONTACT: Mr. Charles Nairn

Center for Rate-Controlled Recordings
University of Louisville
Louisville, Kentucky 40208
CONTACT: Dr. Emerson Foulke

Department of Educational Psychology
School of Education
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin 53201
CONTACT: Dr. Gerald Gleason

HumRRO
Division No. 3
Presidio of Monterey
California 93940
CONTACT: Dr. T.G. Sticht

Speech Research Laboratory & Hearing Clinic
University of Illinois
Urbana, Illinois 61801
CONTACT: Dr. Willard Zemlin

University of Missouri
School of Medicine
Columbia, Missouri 65201
CONTACT: Mrs. Gloria Boyle

Bureau of AV Instruction
National Center for Audio Tapes
University of Colorado
Boulder, Colorado 80302
CONTACT: Mr. Ed Tracy